ehp ENVIRONMENTAL HEALTH PERSPECTIVES
http://www.ehponline.org

Micronuclei in Cord Blood Lymphocytes and Associations with Biomarkers of Exposure to Carcinogens and Hormonally Active Factors, Gene Polymorphisms, and Gene Expression: The NewGeneris Cohort

Domenico Franco Merlo, Silvia Agramunt, Lívia Anna, Harrie Besselink, Maria Botsivali, Nigel J. Brady, Marcello Ceppi, Leda Chatzi, Bowang Chen, Ilse Decordier, Peter B. Farmer, Sarah Fleming, Vincenzo Fontana, Asta Försti, Eleni Fthenou, Fabio Gallo, Panagiotis Georgiadis, Hans Gmuender, Roger W. Godschalk, Berit Granum, Laura J. Hardie, Kari Hemminki, Kevin Hochstenbach, Lisbeth E. Knudsen, Manolis Kogevinas, Katalin Kovács, Soterios A. Kyrtopoulos, Martinus Løvik, Jeanette K. Nielsen, Unni Cecilie Nygaard, Marie Pedersen, Per Rydberg, Bernadette Schoket, Dan Segerbäck, Rajinder Singh, Jordi Sunyer, Margareta Törnqvist, Henk van Loveren, Frederik J. van Schooten, Kim Vande Loock, Hans von Stedingk, John Wright, Jos C. Kleinjans, Micheline Kirsch-Volders, Joost H.M. van Delft, and the NewGeneris Consortium

http://dx.doi.org/10.1289/ehp.1206324

Received: 27 November 2012
Accepted: 18 November 2013
Advance Publication: 19 November 2013

NIH National Institute of Environmental Health Sciences

Micronuclei in Cord Blood Lymphocytes and Associations with Biomarkers of Exposure to Carcinogens and Hormonally Active Factors, Gene Polymorphisms, and Gene Expression: The NewGeneris Cohort

Domenico Franco Merlo,[1] Silvia Agramunt,[2] Lívia Anna,[3] Harrie Besselink,[4] Maria Botsivali,[5] Nigel J. Brady,[6] Marcello Ceppi,[1] Leda Chatzi,[7] Bowang Chen,[8] Ilse Decordier,[9] Peter B. Farmer,[6] Sarah Fleming,[10] Vincenzo Fontana,[1] Asta Försti,[8,11] Eleni Fthenou,[7] Fabio Gallo,[1] Panagiotis Georgiadis,[5] Hans Gmuender,[12] Roger W Godschalk,[13] Berit Granum,[14] Laura J. Hardie,[10] Kari Hemminki,[8,11] Kevin Hochstenbach,[13] Lisbeth E. Knudsen,[15] Manolis Kogevinas,[2,16,17,18] Katalin Kovács,[3] Soterios A. Kyrtopoulos,[5] Martinus Løvik,[14] Jeanette K Nielsen,[15] Unni Cecilie Nygaard,[14] Marie Pedersen,[2,15,16,17] Per Rydberg,[19] Bernadette Schoket,[3] Dan Segerbäck,[20] Rajinder Singh,[6] Jordi Sunyer,[16] Margareta Törnqvist,[19] Henk van Loveren,[13,21] Frederik J. van Schooten,[13] Kim Vande Loock,[9] Hans von Stedingk,[19] John Wright,[21] Jos C. Kleinjans,[13] Micheline Kirsch-Volders,[9] Joost H.M. van Delft,[13] and the NewGeneris Consortium

[1]Epidemiology, Biostatistics, and Clinical Trials, Istituto di Ricerca e Cura a Carattere Scientifico (IRCCS) Azienda Ospedaliera Universitaria (AOU) San Martino-Istituto Nazionale per la Ricerca sul Cancro (IST), Genoa, Italy; [2]Hospital del Mar Medical Research Institute, Barcelona, Spain; [3]Molecular and Environmental Epidemiology, National Institute of Environmental Health, Budapest, Hungary; [4]BioDetection Systems b.v., Amsterdam, Netherlands; [5]National Hellenic Research Foundation, Institute of Biology, Medicinal Chemistry and Biotechnology, Athens, Greece; [6]Cancer Studies and Molecular Medicine, University of

Leicester, Leicester, United Kingdom; [7]Department of Social Medicine, University of Crete, Heraklion, Greece; [8]Division of Molecular Genetic Epidemiology , German Cancer Research Center, Heidelberg, Germany; [9]Laboratory of Cell Genetics, Faculty of Science and Bio-engineering, Vrije Universiteit Brussel, Brussels, Belgium; [10]Division of Biostatistics, Leeds Institute for Genetics, Health and Therapeutics, University of Leeds, Leeds, United Kingdom; [11]Center for Primary Health Care Research, Lund University, Malmö, Sweden; [12]Genedata AG, Basel, Switzerland; [13]Department of Toxicology and Toxicogenomics, Maastricht University, Maastricht, Netherlands; [14]Division of Environmental Medicine, Norwegian Institute of Public Health, Oslo, Norway; [15]Department of Public Health, University of Copenhagen, Copenhagen, Denmark; [16]Centre for Research in Environmental Epidemiology, Barcelona, Spain; [17]CIBER Epidemiología y Salud Pública, Barcelona, Spain; [18]National School of Public Health, Athens, Greece; [19]Environmental Chemistry Unit, Department of Materials and Environmental Chemistry, Stockholm University, Stockholm, Sweden, [20]Department of Biosciences and Nutrition, Karolinska Institute, Huddinge, Sweden; [21]Laboratory for Health Protection Research, National Institute of Public Health and the Environment, Bilthoven, Netherlands.

Corresponding author:

Jos C.S. Kleinjans

Department of Toxicogenomics, Maastricht University

Universiteitssingel 50

6229ER Maastricht

The Netherlands

Email: j.kleinjans@maastrichtuniversity.nl

Internet: www.toxicogenomics-um.nl

Phone: +31433881097

Running title: Micronuclei, exposure biomarkers, and gene expression in newborns

Acknowledgments:

This work was supported by the EU Integrated Project "Newborns and Genotoxic exposure risks; NewGeneris", 6th Framework Programme, Priority 5: Food Quality and Safety [FOOD-CT-2005-016320]. The study was also supported by grants obtained locally including: the National Institute for Health Research, UK (Programme grant RP-PG-0407-10044), the Norwegian Ministry of Health, the Norwegian Ministry of Education and Research, the Norwegian Research Council/FUGE (Grant 151918/S10), and the EU funded HiWATE (Contract Food-CT-2006-036224). Finally, we thank the participants in the study and the doctors, nurses, midwives and laboratory technicians who assisted with its conduct.

Other members of the NewGeneris consortium who participated in this study are: Ricardo Marcos (Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, Bellaterra, Spain), Diana Anderson (Institute for Health Research, Division of Medical Science, University of Bradford, Bradford, United Kingdom), Elena Stagi (Epidemiology, Biostatistics, and Clinical Trials, IRCCS AOU San Martino-IST, Istituto Nazionale per la Ricerca sul Cancro, Genoa, Italy, Viktória Lukács (Molecular and Environmental Epidemiology, National Institute of Environmental Health, Budapest, Hungary), Renee Mijal (Division of Biostatistics, Leeds Institute for Genetics, Health and Therapeutics, University of Leeds, Leeds, United Kingdom), and Ellen Nomark (Division of Environmental Medicine, Norwegian Institute of Public Health, Oslo, Norway).

Competing financial interests declaration: P.R, H.v.S. and M.T. are stakeholders in Adduct Analys AB, Stockholm, Sweden. H.G. is employed by Genedata AG, Basel, Switzerland. H.B.

is employed by BioDetectionSystems, Amsterdam. The other authors and collaborators declare that they have no actual or potential competing financial interests.

ABSTRACT

Background: Leukemia incidence has increased in recent decades amongst European children suggesting early-life environmental exposures play an important role in disease development.

Objectives: We investigated the hypothesis that childhood susceptibility may increase as a result of *in utero* exposure to carcinogens and hormonally acting factors. Utilizing cord blood samples from the NewGeneris cohort, we examined associations between a range of biomarkers of carcinogen exposure and hormonally acting factors with micronuclei (MN) frequency as a proxy measure of cancer risk. Associations with gene expression and genotype were also explored.

Methods: DNA and protein adducts, gene expression profiles, circulating hormonally acting factors, and GWAS data were investigated in relation to genomic damage measured by MN frequency in lymphocytes from 623 newborns enrolled between 2006-2010 across Europe.

Results: Malondialdehyde DNA-adducts (M_1dG) were associated with increased MN frequency in binucleated lymphocytes (MNBN) and exposure to androgenic, estrogenic, and dioxin-like compounds was associated with MN frequency in mononucleated lymphocytes (MNMONO), although no monotonic exposure-outcome relationship was observed. Lower frequencies of MNBN were associated with a 1-unit increase expression of *PDCD11, LATS2, TRIM13, CD28, SMC1A, IL7R, and NIPBL* genes. Gene expression was significantly higher in association with the highest versus lowest category of bulky- and M_1dG-DNA adducts for five and six genes, respectively. Gene expression levels were significantly lower for 11 genes in association with the highest versus lowest category of plasma AR-CALUX® (8 genes), ERα-CALUX® (2 genes), and DR-CALUX®. Several SNPs on chromosome 11 near *FOLH1* significantly modified associations between androgen activity and MNBN frequency. Polymorphisms in *EPHX1/2* and *CYP2E1* were associated with MNBN.

Conclusion: We measured *in utero* exposure to selected environmental carcinogens and circulating hormonally acting factors and detected associations with MN frequency in newborns circulating T-lymphocytes. The results highlight mechanisms that may contribute to carcinogen-induced leukemia and require further research.

INTRODUCTION

Cancer incidence among European children, specifically leukemia, has steadily increased over the last three decades (Kaatsch 2010). In view of the relatively short latent period for leukemia and its very early onset in childhood, it has been suggested that fetal exposure to environmental carcinogens may increase susceptibility to this cancer (Wild and Kleinjans 2003).

The EU funded project "Newborns and Genotoxic exposure risks" (NewGeneris) was designed to evaluate the hypothesis that maternal intake of dietary and other environmental carcinogens results in *in utero* exposure and early biological effects in the unborn child, possibly leading to increased risk of cancer in later childhood (Merlo et al. 2009). The primary aim of the present study was to investigate the relationship between biomarkers of exposure to carcinogenic compounds and micronuclei (MN) frequency in umbilical cord blood lymphocytes from the NewGeneris mother-child birth cohort. The secondary aim was to ascertain whether individual genotypes modify these relationships.

An array of exposure biomarkers quantifying a range of potentially carcinogenic exposures, many of dietary origin, were measured in fetal cord blood samples. DNA and hemoglobin (Hb) adducts reflect biologically effective doses of exposure to genotoxic agents (Phillips, 2008). The DNA-adducts selected for study were 3-(2'-Deoxy-beta-D-erythro-pentofuranosyl)-pyrimido[1,2-alfa]purin-10(3H)-one (M_1dG), O^6-Methyldeoxyguanosine (O^6-MedG), and bulky DNA-adducts. Hb-adducts from acrylamide (AA), glycidamide (GA), and ethylene oxide (EtO) [for which data in this cohort have been published before (Pedersen et al. 2012)] were also evaluated, and three versions of the chemically activated luciferase gene expression (CALUX®) bioassay were used to assess exposure to androgenic, estrogenic and dioxin-like compounds.

Facilitated by the development of microarray technologies, gene expression based biomarkers have been developed and applied for human biomonitoring purposes (McHale et al. 2011; Rager et al. 2011; Ren et al. 2011; van Leeuwen et al. 2008). Gene expression profiling has the potential to identify new biomarkers of exposure that may simultaneously reflect the earliest biological events in disease pathogenesis. Here, we evaluated the expression of 36 genes that were associated with biomarkers of carcinogen exposure by quantitative real-time PCR (qRT-PCR), (Hochstenbach et al. 2012).

MN frequency was assessed as the primary outcome. MN are a potential biomarker of cancer risk, as increased micronucleated binucleated (MNBN) frequencies in T-lymphocytes have been shown to be associated with cancer risk in adults (Bonassi et al. 2007). MN are small extra nuclear bodies arising in dividing cells that are caused by chromosomal breakage and/or whole chromosome loss (Fenech 2007; Kirsch-Volders et al. 2011). MNBN provide a measure of the lesions that have recently occurred *in vivo*, whereas micronucleated mononucleated lymphocytes (MNMONO) give an estimation of the genome damage accumulated over a long period in stem cells and circulating lymphocytes (Kirsch-Volders and Fenech 2001).

Furthermore, we performed a genome-wide-association-study (GWAS) to investigate whether associations between exposure biomarkers and MN are modified by genetic variation.

MATERIALS AND METHODS

Study population and sample collection

Pregnant women (N=1,200) were enrolled between 2006 and 2010 in Heraklion, Greece; Barcelona and Sabadell, Spain; Bradford, England; Copenhagen, Denmark; and in Oslo and

Akerhus, Norway (Pedersen et al. 2012). The participation of mothers in the study was based on previously described eligibility criteria (Pedersen et al. 2012). Study protocols were approved by local ethics committees and informed consent was obtained from all participating mothers prior to sample collection.

Detailed information on personal characteristics, including demographic, health and lifestyle factors, was obtained using extensive questionnaires completed by mothers before or around the time of delivery. Information on dietary habits during pregnancy was obtained from country-specific food-frequency questionnaires (FFQs). Information on birth weight, gestational age, gender, and type of delivery was obtained from maternity records. Gestational age (completed weeks) was computed based on last menstrual period or ultrasound-based estimated date of conception.

Blood samples were collected from 1151 mother-infant dyads following a common protocol as described previously (Merlo et al. 2009). Umbilical cord blood samples were collected immediately after birth from the cord vein of newborns and locally processed. Samples were kept at -20 or -80°C until shipment on dry ice to the study laboratories.

Biomarkers of exposure and early biological effect

DR-CALUX® bioassay

Dioxin-like activity, expressed as AhR-mediated activation of the extractable lipid fraction from plasma, was determined through the DR-CALUX bioassay developed by BioDetection System (Murk et al. 1996). Blood was collected in heparinized tubes and plasma was isolated by centrifugation on the day of collection and frozen at -20^{o}C. One to three ml of cord-blood plasma was used for extraction of lipophilic compounds. The procedure for the DR-CALUX® bioassay

has been described in detail previously (Behnisch 2005). Additional information is provided in Supplemental Material (page 3).

ERα- and AR-CALUX® bioassays

Estrogenic and androgenic activity in cord-blood plasma was determined using the ERα- and AR-CALUX® Bioassays. The ERα and AR-CALUX® bioassays comprise human bone cell lines (U2-OS), stably incorporating the firefly luciferase gene coupled to Responsive Elements (REs) as a reporter gene for the presence of (xeno-) estrogens (Estrogen Responsive, ERα-CALUX®) and androgens (Androgen Responsive, AR-CALUX®) (Sonneveld et al. 2005). Additional information is provided in Supplemental Material (page 4).

Hb-adducts

Erythrocytes were isolated by centrifugation on the day of collection and stored at -20°C. AA-, GA- and EtO-Hb-adducts were simultaneously determined by the "adduct FI*R*E procedure" using liquid chromatography tandem mass spectrometry with performance and validation standards as described in detail elsewhere (von Stedingk et al. 2010; von Stedingk et al. 2011). In total Hb adduct levels were measured in 1151 cord blood samples.

DNA-adducts

DNA was isolated with Qiagen Midi Kit Cat. No. 13343, (Qiagen, Hilden, Germany) with some modifications of the manufacturer's protocol as reported previously (Kovács et al. 2011). Additional details are provided in Supplemental Material (pages 4-8).

Immunoslot blot analysis of M_1dG

M_1dG was determined by an immunoslot blot method, using a murine M_1dG monoclonal primary antibody (D10A1), provided by Prof. Lawrence Marnett (Vanderbilt University, TN, USA), as described previously (Singh et al. 2001).

Immunochemical assays for analyses of O^6-MedG and PAH-DNA-adducts

These analyses were carried out using ultrasensitive sandwich chemiluminescence immunoassays as previously described for O^6-MedG (Georgiadis et al. 2011) and PAH-DNA-adducts (Georgiadis et al. 2012).

Postlabelling analysis of bulky DNA-adducts

Bulky DNA-adducts were detected with the nuclease P1 modification of the ^{32}P-postlabelling procedure as detailed elsewhere (Kovács et al. 2011). Inter-laboratory differences in levels were adjusted for, as described in Supplemental Material (page 8).

Cytokinesis-block micronucleus assay

The *in vitro* cytokinesis blocked MN assay was carried out according to the standardized protocol developed for semi-automated image analysis (Decordier et al. 2009) and adapted for umbilical blood (Vande Loock et al. 2011). MN were scored in both binucleated (MNBN) and mononucleated (MNMONO) T-lymphocytes (Kirsch-Volders and Fenech 2001). To harmonize slide preparation, the cohort cytologists were trained by I. D., K. V.L. and M. K-V (Vrije Universiteit Brussel, VUB). Slides were sent to VUB where staining and MN analysis occurred. Quality control step after staining included visual selection of slides with a good quality using a light microscope and based on a good spreading, swelling and amount of cells. The automated

scoring procedure followed by visual validation of selected micronucleated cells was carried out by the same researcher, using the PathFinder™ platform installed by IMSTAR S.A., Paris, France, at the VUB laboratory consisting of a PathFinder™ CELLSCAN™ capture station and two PathFinder™ MN analysis workstations. Reproducibility of the automated image analysis combined with the visual validation step was investigated by assessing the inter-capturing variability (Decordier et al., 2009; Decordier et al., 2011). At the end of the processing step, cells containing detected MN are presented one by one on the screen and confirmed or rejected by the scorer, according to the Human MicroNucleus project (HUMN) scoring criteria (Fenech et al., 2003). According to guideline T487 of the Organisation for Economic Co-operation and Development (OECD), only subjects with at least 1000 BN lymphocytes counted were considered for statistical analysis (OECD, 2010).

Gene expression analysis

To preserve RNA for gene expression analysis, 0.4 ml of heparin-anticoagulated whole blood was mixed with 1.2 ml of RNAlater (Ambion/Applied Biosystems, Nieuwerkerk aan den Ijssel, the Netherlands) as soon as possible after blood collection. Samples were kept at -80°C until shipment on dry ice to the research laboratory at Maastricht University. Total RNA was isolated using the RiboPure-Blood system (Ambion) according to the manufacturer's instructions. RNA integrity was verified by gel electrophoresis (2100 BioAnalyzer, Agilent Technologies, Amstelveen, the Netherlands).

Fluidigm's BioMark™ Dynamic Array technology was used for gene expression analyses by qRT-PCR, which was conducted by ServiceXS, Leiden, The Netherlands. Thirty-six genes were selected from a whole genome gene-environment interaction study on neonates (n = 84) from the

Norwegian cohort (Hochstenbach et al. 2012). Selection was primarily based on correlations ($r \geq 0.75$; ≤ -0.75, $p < 0.05$) of gene expression with toxic dietary exposures (i.e. genotoxic or immunotoxic) estimated based on food frequency questionnaires, CALUX assay-based evidence of exposure to estrogenic-, androgenic-, and dioxin-like compounds, Hb adduct levels, and MN frequencies. Only mechanistically relevant genes were selected, based on gene ontologies such as "DNA repair", "cell cycle", "apoptosis" and "cell proliferation". For each of the correlations, mechanistically relevant genes were selected, which resulted in total in 36 unique genes. Five reference genes were selected, based on low variance across all individuals. TaqMan gene expression assays (Applied Biosystem, Nieuwekerk aan de Ijssel, the Netherlands) were used (see Supplemental Material, Table S1) and qRT-PCR was conducted according to the manufacturer's protocol. Each sample was analyzed in duplicate, and an average Ct value obtained. On all RT-PCR plates a reference sample at various dilutions was included for quality control assessment of inter-plate reproducibility. The raw Ct value upper cut off was set to 26; genes exceeding this value were classified as unexpressed. For normalization, the average Ct of the five reference genes was subtracted from the Ct value of each gene.

Genome Wide Association Studies and candidate genes analyses

A genome-wide scan of ~300,000 tagging single nucleotide polymorphisms (SNPs) was conducted using the Illumina HumanCytoSNP-12 v1 according to the manufacturer's protocols. Genotype calling was done using Illumina GenomeStudio 2010. Genomic DNA was isolated from 900 cord-blood samples and was used to genotype each child. Quality control was performed on a per-sample and SNP basis. 33 duplicates, 23 samples with a genotype call rate <98.5%, and 14 twins were excluded leaving 830 genotyped samples available for analysis. We used a general genetic model retaining the three distinct genotypes and without making any

assumption about the direction of the SNPs association in the heterozygote compared to the two homozygote classes. According to non-mutually exclusive SNP-based quality checks, 6801 SNPs were excluded because of Hardy-Weinberg equilibrium violation ($p<10^{-6}$), 35429 because they had a minor allele frequency (MAF <1%), and 7338 because missing genotype was >10%, resulting in 258246 out of 298199 SNPs left for statistical analyses. 435 newborns had both SNPs and MN results available and they were used in GWAS statistical analyses.

In addition SNPs present in metabolism and DNA repair genes were selected *a priori* by the consortium as candidate genes based on the available knowledge on functionalities with respect to bioactivation (*CYP1A1, CYP2E1, CYP2D6, EPHX1* and *EPHX2)* and detoxification (*GSTM1)* of DNA adduct-forming metabolites, base excision repair of oxidative adducts (*OGG1),* nucleotide excision repair of bulky adducts *(XRCC1, ERCC2/XPD, XPA,* and *XRCC3),* repair of alkylated adducts (*MGMT, ALKB,* and *MPG)* and of thymine adducts (*TDG*), and with respect to folate metabolism which is known to interfere with micronucleus formation (*MTHFR, MTR* and *MTRR*.)

Statistical analyses

Separate negative binomial fixed effects multivariable regression models were used to estimate the associations of MN frequencies per 1,000 binucleated (MNBN) or mononucleated (MNMONO) T-lymphocytes (as dependent variables) with AA-Hb, GA-Hb, and EtO-Hb adducts (pmol/g Hb, categorized into quintiles); M_1dG-, PAH-, bulky-, and O^6-MedG-DNA adducts (per 10^8 nucleotides, categorized into quartiles); and plasma levels of AR-CALUX® (ng DHT AEQ/ml plasma), ERα-CALUX® (ng 17-®-estradiol EEQ/ml plasma), and DR-CALUX® (pg TEQ/ml plasma) (categorized into quartiles); gene expression (normalized Ct value,

continuous), GWAS data (i.e., all SNPs and *a priori* selected candidate genes), and the interactions between biomarkers and SNPs.

Cohort (country), maternal age (continuous), gestational age (continuous), pre-pregnancy maternal body mass index (continuous), maternal smoking during pregnancy (any or none), environmental tobacco smoke (ETS) exposure during pregnancy (any or none), maternal ethnicity (Caucasian, others), new-born gender and birth weight (continuous), were selected as potential confounders *a priori* and included in all models. Observations with missing covariates were excluded from the statistical analysis. We report the relative difference in the frequency of MN for each category of exposure relative to the lowest exposure category and the associations between 1-unit increases in gene expression and MN frequency as the mean ratio (MR) and its 95% confidence interval (95%CI). The likelihood ratio test was used as a global test of statistical significance over all categories of each exposure biomarker, SNPs allele variants, gene expression and the exposure-SNPs and the gene expression-SNPS interactions.

Associations between the expression of each of the 36 genes evaluated and categorical exposure biomarkers were estimated using separate multivariable linear regression models adjusted for the covariates listed above. The F test was used as a global test of statistical significance over all categories of each exposure biomarker. For each exposure biomarker we report the differences in gene expression associated with the highest versus lowest category of exposure biomarkers.

For gene expression and GWAS analyses the estimated p-values were adjusted to account for multiple comparisons using standard methods (Benjamini and Hochberg 1995; Hochberg 1988; Holm 1979). This criterion was used to identify SNPs associated with MN as main predictors or as effect modifiers of the exposure biomarkers-MN and gene expression-MN associations. No

adjustment was made for p-values estimated from the analyses of *a priori* selected candidate genes. P-values less than 0.05 were considered statistically significant. All associations were examined in newborns with MN assay data available (n=623) and with exposure biomarkers, gene expression, and GWAS data available. Sample sizes for individual association analyses varied as indicated in the results.

Manhattan plots for p-values along chromosome and position were made by the Genetic analysis package (gap) for CRAN R 2.11. Statistical analyses were carried out using Stata S.E. version 10.0 (StataCorp, Texas, USA), R (http://cran.r-project.org), and with Genedata Expressionist 7.0 (Genedata AG, Switzerland).

RESULTS

Levels of biomarkers of exposure, (i.e. Hb- and DNA-adducts, AR-, ERα- and DR- CALUX® activity) detected in newborns are reported in Table 1. The numbers of observations for each biomarker varied reflecting the variable amount of biological specimens collected from cord blood and the assays prioritization adopted (i.e., Hb adducts, DNA adducts and Calux® activity). The largest number of observations was available for AA-Hb adducts (n = 1151) and the smallest for DR- CALUX® (n = 725). For all biomarkers large variations were present (e.g., AA-Hb adducts: median = 14.4 pmol/g Hb, range: 4.4, 124.8; M_1dG-DNA adducts: median = 9.9 per 10^8 nucleotides, range: 0.5, 324.7).

Descriptive statistics for MNBN and MNMONO T-lymphocytes are shown in Table 2 by cohort, and by socio-demographic, reproductive, and life-style factors. Again large inter-individual variations were observed within and between cohorts, with the highest level of MN observed in

Greece (MNBN mean = 1.79 ± 1.50 per 1000 binucleated T-lymphocytes) and the lowest in the United Kingdom (MNBN mean = 0.55 ± 0.74).

None of the global tests of associations across all categories of exposure were statistically significant, and there was no evidence of monotonic dose-response trends with increasing levels of exposure for associations of AA-, GA-, or EtO-Hb adducts (quintiles); PAH-, Bulky-, or O^6-MG-DNA-adducts (quartiles); or DR-CALUX® plasma levels (quartiles) and frequencies of MNBN and MNMONO T-lymphocytes (Table 3). A significant overall association was found between M_1dG levels and the frequency of MNBN lymphocytes, although associations relative to the lowest quartile of M1dG were positive for the 2nd and 3rd quartiles and negative for the highest quartile. ERα-CALUX® plasma levels were significantly associated with the frequency of MNBN and MNMONO lymphocytes and AR-CALUX® with the frequency of MNMONO lymphocytes. No monotonic exposure-outcome association was observed between ERα-CALUX® or AR-CALUX® and MN. For ERα-CALUX® a significant negative association with MNBN was detected for the 2nd quartile, followed by a weak non-significant positive association with the 3rd and 4th quartiles while the associations with MNMONO were negative for the 2nd and 4th quartiles. The strongest associations were detected for AR-CALUX® and MNMONO T-lymphocytes and were positive for the 2nd and 3rd quartiles and negative for the 4th quartile.

One-unit increases in the expression of seven of the 36 genes evaluated (*PDCD11, LATS2, TRIM13, CD28, SMC1A, IL7R,* and *NIPBL*) were associated with significantly lower MNBN frequencies, with MR ranging from 0.81 (95%CI: 0.88, 0.96) for PDCD11 to 0.64 (95%CI: 0.77, 0.97) for NIPBL (Figure 1, Panel A). The frequency of MNMONO was not significantly associated with expression of any of the genes tested (data not shown).

In models with gene expression levels as the dependent variable, expression was significantly higher in association with the highest versus lowest category of bulky DNA adducts and of M_1dG levels for five and six genes, respectively (Figure 1, Panel B). Conversely, expression levels were significantly lower for a total of 11 genes in association with the highest versus lowest category of plasma AR-CALUX® (8 genes), ERα-CALUX® (2 genes), and DR-CALUX® (7 genes) (Figure 1, Panel B). Associations with lower levels of exposure are not reported. Six of the seven genes whose expression was associated with significantly lower MNBN frequency (i.e., all except TRIM13, Figure 1, Panel A) were significantly associated with the highest versus lowest category of at least one exposure biomarker (M_1dG, DR- CALUX®, ERα- CALUX®, or AR-CALUX®, Figure 1, Panel B)

GWAS was carried out on 435 newborns with data available for both SNPs and micronuclei. Confounding by population stratification was assessed (Supplemental Material, Figures S1 and S2) and confirmed that genotype variations occurred between population subgroups (i.e., maternal ethnicity and newborns' country of birth), justifying the need for adjustment in statistical analyses. None of the GWAS SNPs were significant predictors of MNBN frequencies (Supplemental Material, Figure S3). Investigation of the exposure biomarkers-SNPs interactions on the occurrence of MNBN revealed a cluster of significant SNPs (on chromosome 11) for AR-CALUX® modelled as a continuous dependent variable (see Supplemental Material, Figure S4). The four SNPs acting as effect modifiers of the relationship between AR-CALUX® and the frequency of MNBN lymphocytes are given in Supplemental Material, Table S2. The association of these SNPs were reported per unit increase of plasma AR-CALUX® and varied according to the allele variants. For each of the SNPs shown, there was a significant positive association between a 1-unit increase in plasma AR-CALUX® and MNBN frequency among participants

with one homozygous genotype, and a significant negative association with the alternate homozygous genotype (e.g., for rs7131537, MR = 2.54; 95%CI: 1.69, 3.75 for CC and MR = 0.36; 95%CI: 0.21, 0.60 for AA, with a null association among AC heterozygotes compared to an overall estimated association MR = 1.14; 95%CI: 0.88, 1.47, data not shown).

Furthermore, 89 SNPs from the 18 *a priori* selected candidate genes were investigated for association with MN frequencies. SNPs in *EPHX1, EPHX2* and *CYP2E1* were significantly associated (unadjusted overall p-value<0.05) with the frequency of MNBN lymphocytes (Table 4). None of the candidate gene SNPs were significantly associated with the frequency of MNMONO lymphocytes (data not shown).

DISCUSSION

Here, we show that exposure-biomarkers and T-lymphocyte MN levels are measurable in cord blood, that large variations exist for these in the European newborn population, and also that some of the exposure biomarkers are associated with MN levels (as independent variables) and with gene polymorphisms (when the biomarkers are modeled as dependent variables). This suggests that the fetus may be exposed to carcinogenic chemicals *in utero* via the placenta, and that such exposures may be sufficient to exert early biological effects manifested as an increase in the frequency of MNBN, a marker that has been associated with cancer risk in adults (Bonassi et al. 2007). However, our findings should be interpreted with caution given that associations did not show evidence of consistent dose-response relations with increasing levels of exposure.

M_1dG is the major DNA-adduct arising from malondialdehyde, a genotoxic by-product of lipid peroxidation of polyunsaturated fatty acids with a high number of double bonds that also can be

formed during food preparation (Jeong and Swenberg 2005). A significant overall association was detected between M_1dG adduct levels and MNBN frequency though the positive association was limited to the second and third quartiles, with the highest quartile of M_1dG adducts being associated with the lowest MNBN frequency when compared with the lowest quartile. This association indicates recent exposure to malondialdehyde as MNBN formation reflects recent genetic damage that results in micronuclei formation when cell replication is induced in vitro. No association was found between Hb-adducts with MNMONO frequencies; however fetal exposure to compounds detected by ERα-CALUX and AR-CALUX induced significant increases of MNMONO possibly reflecting genetic damage accumulated during fetal development (Kirsch-Volders and Fenech, 2001). The CALUX assays measure estrogenic-, androgenic-, or dioxin-like activities that could result from a variety of compounds or mixtures of compounds. Consequently, associations cannot be attributed to specific exposures. Infant acute leukemia is a frequent childhood cancer and maternal exposure to hormones during pregnancy has been reported as a potential risk in disease occurrence (Pombo-de-Oliveira and Koifman, 2006). A recent review on MN in neonates and children concluded that exposure to environmental pollutants and radiation leads to increased MN; however, no information was provided on possible associations with other biomarkers of exposure and/or early effect, as presented in the current study (Holland et al. 2011).

The reduced number of samples available for the statistical analyses of the relationships between exposure-biomarkers and MN levels is a limitation of the study and may have introduced false-negative findings. Conversely, some of the detected significant associations may have resulted from the multiple comparisons performed increasing the chance of false-positive findings. In addition none of the observed associations followed a dose-response pattern.

We explored the expression of 36 genes by qRT-PCR as potential new biomarkers of toxic exposure. The expression of seven genes was negatively associated with MNBN (none with MNMONO), namely *SMC1A, LATS2, TRIM13, PDCD11, CD28, IL7R* and *NIPBL*. The expression of these particular genes has previously been shown to be affected by one or more genotoxic carcinogens in experimental models (http://ctd.mdibl.org; (Mattingly et al. 2003)). However, as detailed exposure data were absent, we could not further substantiate the involvement of specific chemicals. By using the dedicated TRANSFAC® software for finding transcription factor expression in our transcriptomic data, no transcription factor could be identified that could regulate all these genes. Given that MN are formed during metaphase/anaphase/telophase transition, it was of interest that the majority of the genes identified are involved in progression through the cell cycle, cell division, spindle formation or DNA damage responses. *SMC1A* encodes a protein that is part of the cohesin protein complex and is involved in sister chromatid cohesion during the cell cycle (Bauerschmidt et al. 2011). The tumor suppressor gene *LATS2* encodes a protein that interacts with centrosome proteins and is required for correct spindle formation (Abe et al. 2006). *TRIM13* encodes a kinase involved in many different cellular processes including proliferation and apoptosis (Nakashima 2002). Furthermore, *CD28* and *PDCD11* are involved in apoptosis (Lacana and D'Adamio 1999; Walker et al. 1998). *NIPBL* is required for association of cohesin with chromosomes, for early processing of double-strand breaks and for the DNA damage checkpoint (Oka et al. 2011). For *IL7R* the biological relevance for its association with MNBN remains unclear.

The expression of six of the seven genes associated with MNBN was also associated with the highest versus lowest level of one or more exposure biomarkers (Figure 1). *CD28, IL7R* and *PDCD11A* were associated with the mutagenic DNA-adduct M_1dG. *CD28* and *PDCD11* are

mainly involved in processes linked to genotoxic stress, i.e., apoptosis and cell cycle (Walker et al. 1998, Lacana and D'Adamio 1999). *LATS2* and *SMC1A* were associated with DR-CALUX®, through which compounds that activate the transcription factor AhR, such as PCDDs, PCDFs, dioxin-like PCBs, and PAHs (Pedersen et al. 2010) are measured; many of the latter are genotoxic. Activation of the AhR participates in pathways such as cell cycle regulation, apoptosis and immune responses (Marlowe and Puga 2005). Although *LATS2* and *SMC1A* are not known to be regulated by AhR, both genes are involved in certain sub-processes of the cell cycle (see above). *NIPBL* was associated with AR-CALUX®, which measures compounds with androgenic activity. Like AhR, AR is a transcription factor and regulates the expression of various genes involved in cell cycle control, apoptosis, cell growth, and differentiation (Heisler et al. 1997). Although *NIPBL* is not known to be regulated by AR, it is linked to genotoxic stress related processes and is involved in the cell cycle through its mediating function in sister chromatid cohesion (Watrin et al. 2006).

In summary, associations between gene expression profiles and MN induction reflect the origin of MN: many of the genes are associated with chromosome breakage or loss, and in particular, interference with spindle and chromatid segregation. Their associations with exposure biomarkers support their relevance in relation to genotoxic processes.

The analysis of genetic susceptibility was conducted using GWAS. A strong signal was observed on chromosome 11 for an interaction with AR-CALUX® on MNBN frequency (Supplemental Material, Table S2, Figure S4). The gene closest to this hotspot is *FOLH1* (folate hydrolase 1) and could thus be the genetic factor that affects this relationship. Several pseudogenes were closer, but excluded as their function is unclear. *FOLH1*, also known as PMSA (prostate-specific membrane antigen), is over-expressed in prostate cancer and is negatively regulated by androgen

(Ghosh et al. 2005). Furthermore, a polymorphism in *FOLH1* associated with lower levels of serum folate and hyperhomocysteinemia has been described (Devlin et al. 2000). Low folate is recognized as a risk factor for chromosome instability (Ames 2001) and MN induction (Fenech and Crott 2002). Possibly, an interaction between androgen exposure and a polymorphism that modulates *FOLH1* expression might affect folate levels and therewith modify MNBN frequencies.

GWAS was carried out on 435 newborns with data available for both SNPs and micronuclei. The relatively small sample size is a limitation of the GWAS analysis and is likely to have introduced a risk of false-negative findings due to reduced statistical power to detect the studied associations. To reduce false-positive findings we accounted for multiple comparisons in our primary GWAS analysis, though candidate gene analyses were not adjusted for multiple comparisons. We identified significant associations between *a priori* selected SNPs in *EPHX1, EPHX2* and *CYP2E1* and the frequency of MNBN lymphocytes (Table 4). These SNPs do not affect the protein code, but might be in linkage disequilibrium with causative variants. However, non-causal associations cannot be ruled out, and further clarification is required given inconsistent associations reported between these genes and MN in the literature (Dhillon et al. 2011).

In this study, samples from almost 1200 newborns were collected. Due to limited sample volumes, the number of biomarker measurements varied from 1151 for the AA-Hb-adduct to 623 for MNBN and 435 newborns had data available for both SNPs and micronuclei. For some analyses data were available for a limited number of observations: between 434 and 424 subjects for the associations between MNBN and candidate SNPs, and less than 220 subjects for the interactions SNPs-exposure biomarkers on MNBN frequency. Although we were able to conduct

association studies between individual exposure markers with MNBN, this seriously limited our ability to investigate the interaction between multiple exposure biomarkers and MNBN.

CONCLUSIONS

In conclusion, we demonstrated that gene expression, lymphocyte MN levels, and a variety of biomarkers of environmental (geno)toxic exposure can be measured in newborn cord blood samples, and that there is inter-individual variation in these markers in the European population. Associations of exposure biomarkers and genes (both at the level of gene expression level and genotype) with MN frequencies may help generate new hypotheses about mechanisms of carcinogen-induced leukemia. The associations that we report must be interpreted with caution as we did not measure specific exposures, we did not observe monotonic dose-response relations, and we cannot rule out non-causal associations.

Nevertheless our results suggest that internal exposure of the fetus to toxic chemicals occurs during apparently normal pregnancies, that such exposures may increase the frequency of MN formation [which, while of uncertain relevance in newborns (Holland et al. 2011) has been associated with cancer risk in adults (Bonassi et al. 2007)], and that some children may be more susceptible to genotoxic effects of *in utero* exposures than others.

Ultimately, information on the effects and sources of *in utero* genotoxic exposures could be used by regulators and industry to develop policy measures and strategies to reduce such exposures in order to improve children's health and reduce the incidence of childhood cancer.

REFERENCES

Abe Y, Ohsugi M, Haraguchi K, Fujimoto J, Yamamoto T. 2006. LATS2-Ajuba complex regulates gamma-tubulin recruitment to centrosomes and spindle organization during mitosis. FEBS Lett 580**:**782-8.

Ames BN. 2001. DNA damage from micronutrient deficiencies is likely to be a major cause of cancer. Mutat Res 475**:**7-20.

Bauerschmidt C, Woodcock M, Stevens DL, Hill MA, Rothkamm K, Helleday T. 2011. Cohesin phosphorylation and mobility of SMC1 at ionizing radiation-induced DNA double-strand breaks in human cells. Exp Cell Res 317**:**330-7.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate - a practical and powerful approach to multiple testing. J R Stat. Soc Se. B-Methodol 57**:**289-300.

Bonassi S, Znaor A, Ceppi M, Lando C, Chang WP, Holland N et al. 2007. An increased micronucleus frequency in peripheral blood lymphocytes predicts the risk of cancer in humans. Carcinogenesis 28**:**625-31.

Decordier I, Papine A, Plas G, Roesems S, Vande Loock K, Moreno-Palomo J et al. 2009. Automated image analysis of cytokinesis-blocked micronuclei: an adapted protocol and a validated scoring procedure for biomonitoring. Mutagenesis 24**:**85-93.

Decordier I, Papine A, Vande Loock K, Plas G, Soussaline F, Kirsch-Volders M. 2011. Automated image analysis of micronuclei by IMSTAR for biomonitoring. 26**:**163-168.

Devlin AM, Ling EH, Peerson JM, Fernando S, Clarke R, Smith AD et al. 2000. Glutamate carboxypeptidase II: a polymorphism associated with lower levels of serum folate and hyperhomocysteinemia. Hum Mol Genet 9**:**2837-2844.

Dhillon VS, Thomas P, Iarmarcovai G, Kirsch-Volders M, Bonassi S, Fenech M. 2011. Genetic polymorphisms of genes involved in DNA repair and metabolism influence micronucleus frequencies in human peripheral blood lymphocytes. Mutagenesis 26**:**33-42.

Fenech M, Chang WP, Kirsch-Volders M, Holland N, Bonassi S, Zeiger E (2003) HUMN project: detailed description of the scoring criteria for the cytokinesis-block micronucleus assay using isolated human lymphocyte cultures. Mutation research 534**:** 65-75

Fenech M. 2007. Cytokinesis-block micronucleus cytome assay. Nat Protoc 2**:**1084-1104

Fenech M, Crott JW. 2002. Micronuclei, nucleoplasmic bridges and nuclear buds induced in folic acid deficient human lymphocytes-evidence for breakage-fusion-bridge cycles in the cytokinesis-block micronucleus assay. Mutat Res 504**:**131-6.

Georgiadis P, Kaila S, Makedonopoulou P, Fthenou E, Chatzi L, Pletsa V et al. 2011. Development and validation of a new, sensitive immunochemical assay for O(6)-methylguanine in DNA and its application in a population study. Cancer Epid Biom Prev 20**:**82-90.

Georgiadis P, Kovacs K, Kaila S, Makedonopoulou P, Anna L, Poirier MC et al. 2012. Development and validation of a direct sandwich chemiluminescence immunoassay for measuring DNA adducts of benzo[a]pyrene and other polycyclic aromatic hydrocarbons. Mutagenesis 27**:**589-597.

Ghosh A, Wang X, Klein E, Heston WD. 2005. Novel role of prostate-specific membrane antigen in suppressing prostate cancer invasiveness. Cancer Res 65**:**727-31.

Heisler LE, Evangelou A, Lew AM, Trachtenberg J, Elsholtz HP, Brown TJ. 1997. Androgen-dependent cell cycle arrest and apoptotic death in PC-3 prostatic cell cultures expressing a full-length human androgen receptor. Mol Cell Endocrinol 126**:**59-73.

Hochberg Y. 1988. A sharper Bonferroni procedure for multiple tests of significance. Biometrika 75**:**800-803.

Hochstenbach K, van Leeuwen DM, Gmuender H, Gottschalk RW, Løvik M, Granum B et al. 2012. Global Gene Expression Analysis in Cord Blood Reveals Gender-Specific Differences in Response to Carcinogenic Exposure In Utero. Cancer Epid Biom Prev 10**:**1756-1767.

Holland N, Fucic A, Merlo DF, Sram R, Kirsch-Volders M. 2011. Micronuclei in neonates and children: effects of environmental, genetic, demographic and disease variables. Mutagenesis 26**:**51-6

Holm S. 1979. A simple sequentially rejective multiple test procedure. Scand. J. Statist. 6**:**65-70.

Jeong YC, Swenberg JA. 2005. Formation of M1G-dR from endogenous and exogenous ROS-inducing chemicals. Free Radic Biol Med 39**:**1021-9.

Kaatsch P. 2010. Epidemiology of childhood cancer. Cancer Treat Rev 36**:**277-85.

Kirsch-Volders M, Fenech M. 2001. Inclusion of micronuclei in non-divided mononuclear lymphocytes and necrosis/apoptosis may provide a more comprehensive cytokinesis block micronucleus assay for biomonitoring purposes. Mutagenesis 16**:**51-8.

Kirsch-Volders M, Plas G, Elhajouji A, Lukamowicz M, Gonzalez L, Vande Loock K et al. 2011. The in vitro MN assay in 2011: origin and fate, biological significance, protocols, high throughput methodologies and toxicological relevance. Arch Toxicol 85**:**873-99.

Kovács K, Anna L, Rudnai P, Schoket B. 2011. Recovery of bulky DNA adducts by the regular and a modified 32P-postlabelling assay; influence of the DNA-isolation method. Mutat Res 721**:**95-100

Lacana E, D'Adamio L. 1999. Regulation of Fas ligand expression and cell death by apoptosis-linked gene 4. Nat Med 5**:**542-7.

Marlowe JL, Puga A. 2005. Aryl hydrocarbon receptor, cell cycle regulation, toxicity, and tumorigenesis. J Cell Biochem 96**:**1174-84.

Mattingly CJ, Colby GT, Forrest JN, Boyer JL. 2003. The Comparative Toxicogenomics Database (CTD). Environ Health Perspect 111**:**793-5.

McHale CM, Zhang L, Lan Q, Vermeulen R, Li G, Hubbard AE et al. 2011. Global gene expression profiling of a population exposed to a range of benzene levels. Environ Health Perspect 119**:**628-34.

Merlo DF, Wild CP, Kogevinas M, Kyrtopoulos S, Kleinjans J, NewGeneris C. 2009. NewGeneris: a European study on maternal diet during pregnancy and child health. Cancer Epid Biom Prev 18**:**5-10.

Murk AJ, Legler J, Denison MS, Giesy JP, van de Guchte C, Brouwer A. 1996. Chemical-activated luciferase gene expression (CALUX): a novel in vitro bioassay for Ah receptor active compounds in sediments and pore water. Fundam Appl Toxicol 33**:**149-60.

Nakashima S. 2002. Protein kinase C alpha (PKC alpha): regulation and biological function. J Biochem 132**:**669-75.

OECD-Organisation for Economic Co-operation and Development. 2010. Guideline for Testing of Chemicals. No. 487: In Vitro Mammalian Cell Micronucleus Test (MNvit). OECD, Paris (available at: www.oecd.org/env/testguidelines.

Oka Y, Suzuki K, Yamauchi M, Mitsutake N, Yamashita S. 2011. Recruitment of the cohesin loading factor NIPBL to DNA double-strand breaks depends on MDC1, RNF168 and HP1gamma in human cells. Biochem Biophys Res Commun 411**:**762-7.

Pedersen M, Halldorsson TI, Mathiesen L, Mose T, Brouwer A, Hedegaard M et al. 2010. Dioxin-like exposures and effects on estrogenic and androgenic exposures and micronuclei frequency in mother-newborn pairs. Environ Int 36**:**344-51.

Pedersen M, von Stedingk H, Botsivali M, Agramunt S, Alexander J, Brunborg G et al. 2012. Birth Weight, Head Circumference, and Prenatal Exposure to Acrylamide from Maternal Diet: The European Prospective Mother-Child Study (NewGeneris). Environ Health Perspect 12**:**1739-1745.

Phillips DH. 2008. Biomarkers of exposure. In: Molecular Epidemiology of Chronic Diseases (Wild C, Vineis P, Garte S, eds), Chichester: John Wiley & Sons Ltd.

Pombo-de-Oliveira MS, Koifman S. 2006. Infant acute leukemia and maternal exposures during pregnancy. Cancer Epid Biom Prev 15:2336-2341.

Rager JE, Smeester L, Jaspers I, Sexton KG, Fry RC. 2011. Epigenetic changes induced by air toxics: formaldehyde exposure alters miRNA expression profiles in human lung cells. Environ Health Perspect 119**:**494-500.

Ren X, McHale CM, Skibola CF, Smith AH, Smith MT, Zhang L et al. 2011. An emerging role for epigenetic dysregulation in arsenic toxicity and carcinogenesis. Environ Health Perspect 119**:**11-9.

Singh R, Leuratti C, Josyula S, Sipowicz MA, Diwan BA, Kasprzak KS et al. 2001. Lobe-specific increases in malondialdehyde DNA adduct formation in the livers of mice following infection with Helicobacter hepaticus. Carcinogenesis 22**:**1281-7.

Sonneveld E, Jansen HJ, Riteco JA, Brouwer A, van der Burg B. 2005. Development of androgen- and estrogen-responsive bioassays, members of a panel of human cell line-based highly selective steroid-responsive bioassays. Toxicol Sci 83**:**136-48.

van Leeuwen DM, Pedersen M, Hendriksen PJ, Boorsma A, van Herwijnen MH, Gottschalk RW et al. 2008. Genomic analysis suggests higher susceptibility of children to air pollution. Carcinogenesis 29**:**977-83.

Vande Loock K, Fthenou E, Decordier I, Chalkiadaki G, Keramarou M, Plas G et al. 2011. Maternal and gestational factors and micronucleus frequencies in umbilical blood: the NewGeneris Rhea cohort in Crete. Environ Health Perspect 119**:**1460-5.

von Stedingk H, Rydberg P, Tornqvist M. 2010. A new modified Edman procedure for analysis of N-terminal valine adducts in hemoglobin by LC-MS/MS. J Chrom B Analyt Technol Biomed Life Sci 878**:**2483-90.

von Stedingk H, Vikstrom AC, Rydberg P, Pedersen M, Nielsen JK, Segerback D et al. 2011. Analysis of hemoglobin adducts from acrylamide, glycidamide, and ethylene oxide in paired mother/cord blood samples from Denmark. Chem Res Toxicol 24**:**1957-65.

Walker LS, McLeod JD, Boulougouris G, Patel YI, Hall ND, Sansom DM. 1998. Down-regulation of CD28 via Fas (CD95): influence of CD28 on T-cell apoptosis. Immunology 94**:**41-7.

Watrin E, Schleiffer A, Tanaka K, Eisenhaber F, Nasmyth K, Peters JM. 2006. Human Scc4 is required for cohesin binding to chromatin, sister-chromatid cohesion, and mitotic progression. Curr Biol 16**:**863-74.

Wild CP, Kleinjans J. 2003. Children and increased susceptibility to environmental carcinogens: evidence or empathy? Cancer Epid Biom Prev 12**:**1389-94.

Table 1. Biomarkers of exposure measured in cord blood: descriptive statistics.

Biomarker of exposure (unit)	N	Mean ± SD	Median	IQR	Min, Max
Acrylamide-Hb-adducts (pmol/g Hb)	1151	19.7 ± 16.5	14.4	10.8, 21.6	4.4, 124.8
Glycidamide -Hb-adducts (pmol/g Hb)	1150	13.7 ± 10.4	10.8	7.9, 15.7	2.0, 103.2
Ethylene oxide-Hb-adducts (pmol/g Hb)	1123	13.3 ± 13.8	9.4	6.6, 14.5	0.5, 174.2
M_1dG (per 10^8nucleotides)	892	53.1 ± 52.5	34.2	12.2, 82.7	0.5, 324.7
PAH-DNA-adducts (per 10^8 nucleotides)	797	1.89 ± 2.09	1.39	0.58, 2.55	0.2, 28.3
Bulky DNA-adducts (per 10^8 nucleotides)	635	11.8 ± 11.2	8.4	4.8, 15.3	0.6, 116.6
O^6-MedG (per 10^8 nucleotides)	865	0.470 ± 0.501	0.287	0.075, 0.665	0.015, 3.033
AR-CALUX® (ng DHT AEQ/ml plasma)	765	0.108 ± 0.098	0.099	0.059, 0.140	0.015, 2.142
ERα-CALUX® (ng 17-β-estradiol EEQ/ml plasma)	765	22.1 ± 19.4	18.2	9.5, 28.5	0.01, 182.7
DR-CALUX® (pg TEQ /ml plasma)	725	0.165 ± 0.137	0.130	0.055, 0.230	0.055, 1.043

SD: standard deviation; IQR: inter-quartile range.

Table 2. MNBN and MNMONO T-lymphocytes measured in cord blood (No.= 623) by socio-demographic, reproductive and life-style factors: descriptive statistics.

Covariates	N (%)	MNBN[a] Mean±SD	MNMONO[a] Mean±SD
Country			
United Kingdom	143 (22.9)	0.55 ± 0.74	0.04 ± 0.14
Greece	232 (37.2)	1.79 ± 1.50	0.62 ± 0.71
Denmark	142 (22.7)	0.70 ± 0.58	0.17 ± 0.58
Spain	70 (11.2)	1.00 ± 1.00	0.20 ± 0.45
Norway	36 (5.77)	1.16 ± 0.92	0.11 ± 0.42
Maternal age (years)			
≤ 27	170 (27.2)	1.31 ± 1.26	0.34 ± 0.59
28-30	131 (21.0)	1.11 ± 1.36	0.27 ± 0.53
31-32	96 (15.4)	1.11 ± 1.09	0.34 ± 0.61
33-35	109 (17.4)	1.17 ± 1.34	0.37 ± 0.73
≥ 36	112 (17.9)	0.85 ± 0.84	0.21 ± 0.55
Unknown	5 (0.80)	1.15 ± 1.43	0.20 ± 0.38
Pre-pregnancy BMI			
≤ 18.5 (underweight)	27 (4.33)	1.44 ± 1.19	0.39 ± 0.51
18.6-25.0 (normal)	316 (50.7)	1.20 ± 1.19	0.33 ± 0.65
25.1-30.0 (overweight)	112 (17.9)	1.27±1.36	0.25 ± 0.44
>30 (obese)	78 (12.5)	1.07 ± 1.28	0.47 ± 0.78
Unknown	90 (14.4)	0.68 ± 0.92	0.13 ± 0.32
Birth weight			
Normal (≥ 2500 g)	601 (96.4)	1.13 ± 1.22	0.31 ± 0.60
Low (< 2500 g)	17 (2.72)	1.24 ± 0.88	0.37 ± 0.49
Unknown	5 (0.80)	1.32 ± 1.83	0.19 ± 0.38
Child gender			
Boys	315 (50.5)	1.10 ± 1.22	0.32 ± 0.63
Girls	306 (49.1)	1.16 ± 1.21	0.30 ± 0.57
Unknown	2 (0.32)	0.37 ± 0.07	0.00 ± 0.00
Maternal ethnicity			
Caucasian	522 (83.7)	1.21 ± 1.26	0.35 ± 0.64
Others	93 (14.9)	0.72 ± 0.86	0.08 ± 0.28
Unknown	8 (1.28)	0.44 ± 0.45	0.00 ± 0.00
Gestational age (weeks)			
<37 weeks	29 (4.65)	1.32 ± 1.62	0.30 ± 0.38
≥ 37 weeks	592 (95.0)	1.12 ± 1.19	0.31 ± 0.61
Unknown	2 (0.32)	1.22 ± 1.13	0.00 ± 0.00
Delivery			
Vaginal	287 (46.0)	1.24 ± 1.26	0.35 ± 0.67
Caesarean	334 (53.6)	1.04 ± 1.17	0.27 ± 0.54
Unknown	2 (0.32)	0.48 ± 0.08	0.00 ± 0.00
Maternal smoking during pregnancy			
None	470 (75.4)	1.06 ± 1.16	0.27 ± 0.58
Any	136 (21.8)	1.39 ± 1.37	0.47 ± 0.66
Unknown	17 (2.72)	0.89 ± 1.07	0.09 ± 0.25

Covariates	N (%)	MNBN[a] Mean±SD	MNMONO[a] Mean±SD
ETS exposure during pregnancy			
None	285 (45.7)	0.91 ± 0.91	0.18 ± 0.48
Any	231 (37.0)	1.26 ± 1.43	0.39 ± 0.68
Unknown	107 (17.1)	1.45 ± 1.32	0.46 ± 0.65

[a]per 1000 binucleated or mononucleated T-lymphocytes; ETS: environmental tobacco smoke

Table 3. Relative difference in the frequency of MNBN and MNMONO T-lymphocytes (mean ratio, MR[a]) by category of exposure biomarkers estimated by negative binomial regression.

Exposure biomarker (unit) and category	N (%)	MNBN[b] MR (95%CI)	p-value[c]	MNMONO[b] MR (95%CI)	p-value[c]
Acrylamide-Hb-adducts (pmol/g Hb)					
≤10.2	95 (20.3)	1		1	
10.3-13.9	93 (19.9)	1.06 (0.83, 1.35)		0.79 (0.48, 1.30)	
14.0-18.4	94 (20.1)	1.00 (0.78, 1.28)		0.98 (0.62, 1.55)	
18.5-27.8	93 (19.9)	0.97 (0.74, 1.27)		0.87 (0.53, 1.43)	
>27.8	92 (19.7)	1.24 (0.92, 1.68)		0.83 (0.46, 1.48)	
Total	467 (100)		0.468		0.873
Glycidamide-Hb-adducts (pmol/g Hb)					
≤7.4	91 (19.4)	1		1	
7.5-9.5	97 (20.7)	1.05 (0.83, 1.35)		1.17 (0.71, 1.94)	
9.6-13.1	93 (19.9)	1.09 (0.85, 1.40)		1.10 (0.67, 1.82)	
13.2-18.5	95 (20.3)	0.77 (0.58, 1.02)		0.74 (0.42, 1.31)	
>18.5	91 (19.4)	1.03 (0.76, 1.40)		0.79 (0.44, 1.41)	
Total	467 (100)		0.079		0.429
Ethylene oxide-Hb-adducts (pmol/g Hb)					
≤6.3	89 (19.4)	1		1	
6.4-8.5	94 (20.5)	1.22 (0.94, 1.58)		0.95 (0.55, 1.65)	
8.6-11.5	93 (20.3)	1.22 (0.94, 1.59)		1.12 (0.64, 1.96)	
11.6-16.8	91 (19.9)	1.08 (0.83, 1.41)		1.19 (0.70, 2.03)	
>16.8	90 (19.6)	1.01 (0.76, 1.33)		1.39 (0.81, 2.38)	
Total	457 (100)		0.336		0.676
PAH-DNA-adducts (per 10^8 nucleotides)					
≤6.3	85 (24.5)	1		1	
6.4-8.5	87 (25.1)	0.92 (0.71, 1.18)		0.97 (0.62, 1.53)	
8.6-11.5	87 (25.1)	0.74 (0.56, 0.97)		1.20 (0.71, 2.01)	
11.6-16.8	87 (25.1)	0.98 (0.76, 1.28)		0.89 (0.57, 1.41)	
Total	346 (100)		0.133		0.724
Bulky DNA-adducts (per 10^8 nucleotides)					
≤4.60	77 (25.4)	1		1	
4.70–7.90	73 (24.0)	1.20 (0.91, 1.59)		0.75 (0.37, 1.54)	
8.00–14.78	77 (25.4)	0.90 (0.68, 1.20)		0.78 (0.41, 1.49)	
>14.78	76 (25.0)	0.85 (0.64, 1.14)		0.72 (0.40, 1.30)	
Total	303 (100)		0.096		0.757
M_1dG (per 10^8 nucleotides)					
≤12.02	94 (24.6)	1		1	
12.03–37.78	95 (24.9)	1.16 (0.91, 1.49)		0.99 (0.66, 1.49)	
37.79–83.75	96 (25.1)	1.29 (0.99, 1.67)		1.18 (0.74, 1.88)	
>83.75	96 (25.1)	0.89 (0.69, 1.15)		0.90 (0.58, 1.41)	
Total	381 (100)		0.024		0.761
O^6-MedG (per 10^8 nucleotides)					
≤ 0.076	133 (35.5)	1		1	
0.077-0.389	79 (21.1)	0.99 (0.80, 1.23)		0.88 (0.60, 1.29)	
0.390-0.810	81 (21.6)	1.16 (0.93, 1.45)		1.08 (0.73, 1.61)	

Exposure biomarker (unit) and category	N (%)	MNBN[b] MR (95%CI)	p-value[c]	MNMONO[b] MR (95%CI)	p-value[c]
>0.810	81 (21.6)	1.06 (0.85, 1.34)		1.50 (1.00, 2.24)	
Total	374 (100)		0.539		0.111
AR-CALUX® (ng DHT AEQ/ml plasma)					
≤0.086	65 (25.0)	1		1	
0.087-0.118	63 (24.3)	1.46 (1.07, 2.00)		2.17 (0.99, 4.74)	
0.119-0.151	67 (25.7)	1.37 (0.99, 1.92)		2.63 (1.19, 5.78)	
>0.151	65 (25.0)	1.38 (0.97, 1.96)		0.78 (0.28, 2.17)	
Total	260 (100)		0.103		0.002
ERα-CALUX® (ng 17-β-estradiol EEQ/ml plasma)					
≤11.61	64 (24.8)	1		1	
11.62-18.41	64 (24.8)	0.66 (0.46, 0.95)		0.54 (0.25, 1.20)	
18.42-30.00	65 (25.1)	1.04 (0.77, 1.42)		1.41 (0.74, 2.70)	
>30.00	65 (25.1)	1.09 (0.78, 1.51)		0.57 (0.22, 1.51)	
Total	258 (100)		0.018		0.014
DR-CALUX® (pg TEQ/ml plasma)					
≤0.055	73 (29.6)	1		1	
0.056-0.180	49 (19.9)	0.79 (0.55, 1.12)		1.50 (0.81, 2.77)	
0.181-0.280	63 (25.6)	1.14 (0.85, 1.54)		1.10 (0.57, 2.11)	
>0.281	61 (24.7)	1.12 (0.78, 1.61)		1.74 (0.83, 3.65)	
Total	246 (100)		0.182		0.355

[a]MR = mean ratio adjusted for country, maternal age, pre-pregnancy BMI, birth weight, gender, maternal ethnicity, gestational age, delivery, maternal smoking, and ETS; [b]per 1000 binucleated or mononucleated T-lymphocytes; [c]Log likelihood Ratio Test p-value

Table 4. Relationships between the available SNPs for the *a priori* selected *EPHX* and *CYP2E1* genes and frequency of MNBN T-lymphocytes in newborns.

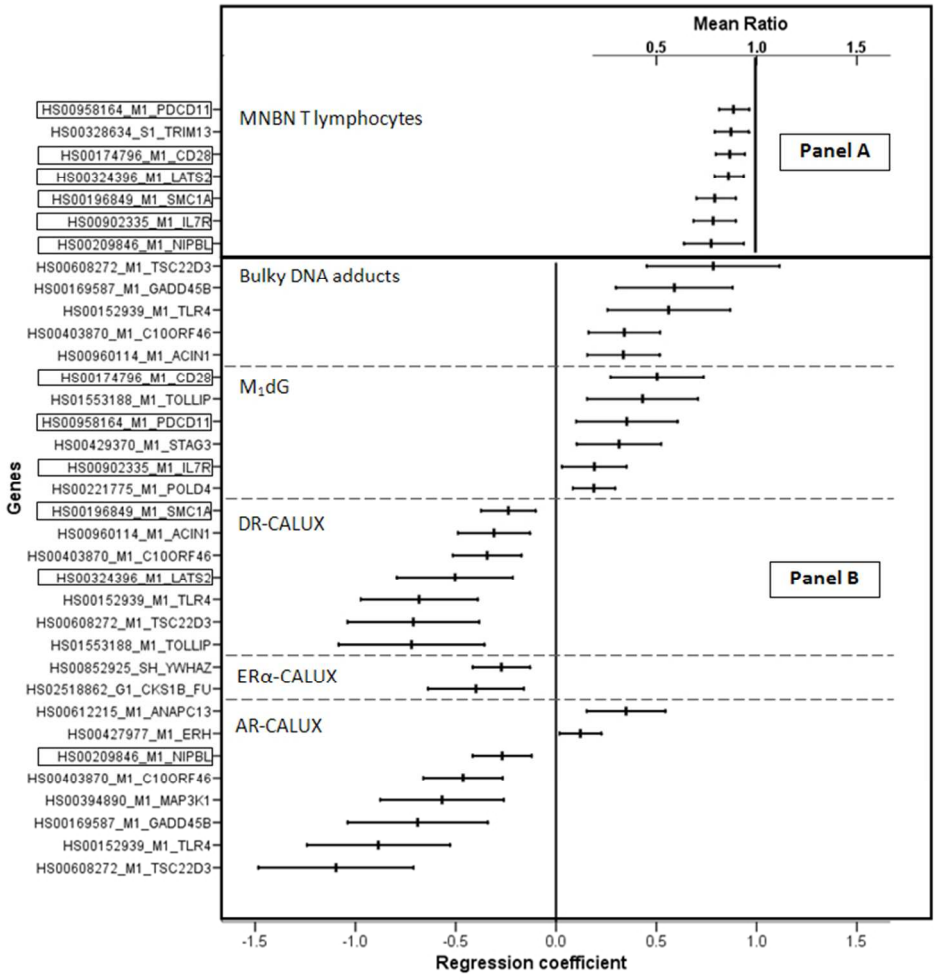
SNP	Gene	N	MNBN[a] Mean (SE)	MR (95CI%)[b]	p-value[c]
Rs1051741	*EPHX1*	424			0.011
GG		330	1.20 (0.07)	1	
AG		85	1.12 (0.11)	1.04 (0.85, 1.27)	
AA		9	0.32 (0.15)	0.35 (0.17, 0.71)	
Rs4149244	*EPHX2*	434			0.032
GG		377	1.17 (0.06)	1	
AG		50	1.04 (0.16)	0.85 (0.66, 1.08)	
AA		7	1.86 (0.48)	1.89 (1.05, 3.39)	
Rs2480258/Rs915906	*CYP2E1*	347			0.040
Other		331	1.03 (0.06)	1	
AA/GG		16	1.14 (0.33)	1.53 (1.02, 2.28)	
Rs2480258	*CYP2E1*	347			0.268
GG		203	1.02 (0.08)	1	
AG		119	1.05 (0.10)	1.02 (0.86, 1.22)	
AA		25	1.08 (0.25)	1.38 (0.96, 1.98)	
Rs915906	*CYP2E1*	347			0.265
AA		228	1.04 (0.08)	1	
AG		102	1.01 (0.10)	1.07 (0.88, 1.29)	
GG		17	1.09 (0.32)	1.50 (0.99, 2.24)	

Analyses carried out on 18 candidate genes with 89 SNPs available on the array. Only statistically significant relationships are reported. For the Rs2480258/Rs915906 SNPs relationships estimated for the single SNPs are shown.

[a]Mean(SE) per 1000 binucleated T-lymphocytes; [b]MR=mean ratio adjusted for country, maternal age, pre-pregnancy BMI, birth weight, gender, maternal ethnicity, gestational age, delivery, maternal smoking, and ETS; [c] Log-Likelihood Ratio Test p-value unadjusted for multiple comparisons.

FIGURE LEGEND

Figure 1. Associations between gene expression and MNBN T-lymphocyte frequency (Panel A) and between exposure biomarkers and gene expression (Panel B) adjusted for country, maternal age, pre-pregnancy BMI, birth weight, gender, maternal ethnicity, gestational age, delivery, maternal smoking, and ETS. Associations shown are those with multiple comparisons-adjusted p-values <0.05. Panel A: mean ratios (MR) for associations between 1-unit increases in gene expression and MNBN based on 350 observations with complete data; Panel B: differences in gene expression associated with the highest versus lowest category of exposure biomarkers based on observations with complete data (bulky DNA adducts n = 398, M_1dG DNA adducts n = 533, AR-CALUX® n = 457; DR-CALUX® n = 477; ER-CALUX® n = 457). DR-CALUX® categories were based on the following centiles: ≤ 0.13; 0.131-0.23; >0.23 (pg TEQ/ml plasma) because 46% of the observations were tied (i.e., below the limit of detection of the assay). Boxed genes are significant predictors of MNBN and are significantly predicted by the highest versus lowest exposure category of at least one exposure biomarker. Panel A: ▌ = mean ratio point estimates of the associations between 1-unit increases in gene expression and MNBN; Panel B: ▌ = regression coefficient point estimates of the differences in gene expression associated with the highest versus lowest category of exposure biomarkers; Horizontal bars = 95%CI.



170x171mm (120 x 120 DPI)